CS 540 Introduction to Artificial Intelligence
Linear Algebra

University of Wisconsin–Madison
Fall 2025, Section 3
September 10, 2025

Announcements

- HW 1 to be released Friday (9/12)
- Optional Review Sessions
  - Held by TA Guy Zamir in Morgridge Hall 3610
  - This Thursday, 5:30-6:30 pm
  - From next week: Tuesdays at 5:30-6:30 pm

Announcements

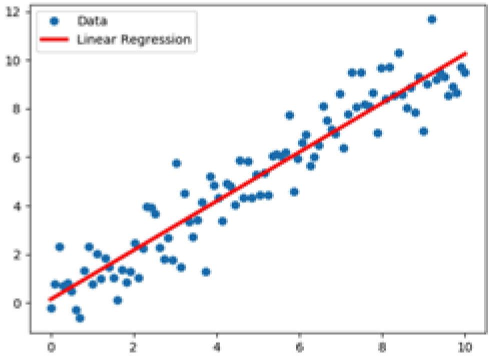
- Class roadmap:

Probability
Statistics
Linear Algebra
Principal Component Analysis (PCA)
Logic
NLP

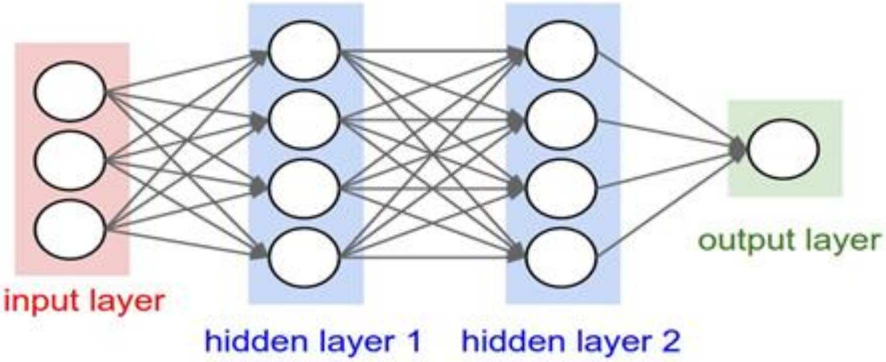
Mostly Foundations

Artificial Intelligence Runs on Linear Algebra

- Study of **linear** functions: simple, tractable
- In AI/ML: building blocks for all models
  - e.g., linear regression; part of neural networks



Hieu Tran



Stanford CS231n

Artificial Intelligence Runs on Linear Algebra

Graphics Processing Units (GPUs) excel at linear algebra workloads

nvidia.com

engineering.fb.com

Today: Linear Algebra Overview

- Basics
- Matrix Multiplication
- Advanced Topics

Mathematical tools for:

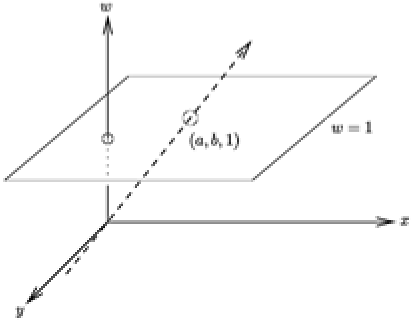
1. Organizing and manipulating high-dimensional data
2. Working with linear functions

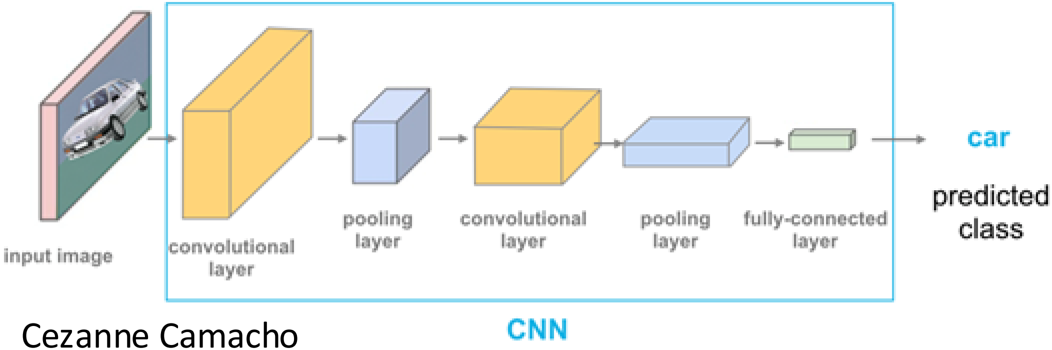
Basics

Basics: **Vectors**

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} \in \mathbb{R}^5$$

- Many interpretations
  - List of values (represents information)
  - Point in a space
- Dimension: number of values: $x \in \mathbb{R}^d$
- AI/ML: often use **very high dimensions:**
  - Ex: images!





Cezanne Camacho

Basics: Matrices

- Many interpretations
  - Array of values
  - List of vectors
  - Description of **linear transformation**

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$$

Basics: Dimensions and Indexing

- Always: (# rows) x (# columns)
- Applies to vectors and scalars
- Indexing: same convention

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$$

$$A \in \mathbb{R}^{4\times 3}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$x \in \mathbb{R}^{2\times 1} = \mathbb{R}^2$$

$$z = [z]$$

$$x \in \mathbb{R}^{1\times 1} = \mathbb{R}$$

$$B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$B_{21} = 3$$

Basics: Transposition

- Transpose: flip rows and columns

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{bmatrix}$$

- Same for vectors
  - Standard vector is row
  - Transpose yields column

Basics: **Addition**

- Entry-by-entry

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1+3 \\ 2+4 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

Matrix Multiplication

Gameplan for Matrix Multiplication

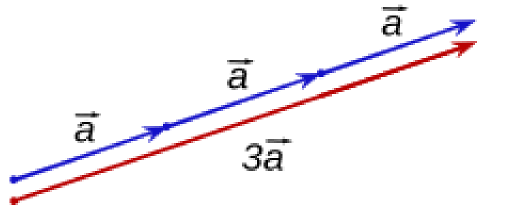
- Addition for matrices and vectors is simple.
- What about multiplication?
- Answer is counterintuitive!
- We will see:
  - Two simple cases
  - The general rule
  - More simple cases

Matrix Multiplication: Vector & Scalar

- How to multiply a vector and a scalar?

$$cx = c\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ cx_3 \end{bmatrix}$$

- Uniform stretch/scaling on each component



Wikipedia

Matrix Multiplication: Diagonal Matrix

- Recall: a matrix represents a linear transformation
- **Diagonal** matrices are simple transformations
- Stretch differently along different directions

$$D = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$Dx = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} d_{11}x_1 \\ d_{22}x_2 \\ d_{33}x_3 \end{bmatrix}$$

Scale different components separately!

Matrix Multiplication in General

- Rule: "middle" dimensions must match
  - To multiply $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{p \times q}$, need $m = p$
  - Yields $C \in \mathbb{R}^{n \times p}$
- Computation for $C = AB$ is an *inner sum*

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \end{bmatrix}$$

Matrix Multiplication: Matrix-Vector

- Multiply
  - $A \in \mathbb{R}^{n \times p}$
  - $x \in \mathbb{R}^{p} = \mathbb{R}^{p \times 1}$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(4×1) (4×3) (3×1)

these must match

result is in $\mathbb{R}^{4 \times 1}$

(4×1)

$$y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3$$

Matrix Multiplication: Outer Product

- Multiply vectors, get a matrix

$$xy^T = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} y_1 & y_2 \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 \\ x_2 y_1 & x_2 y_2 \end{bmatrix}$$

$$x \in \mathbb{R}^2 = \mathbb{R}^{2\times 1}$$

$$y \in \mathbb{R}^2 = \mathbb{R}^{2\times 1}$$

$$y^T \in \mathbb{R}^{1\times 2}$$

$$xy^T \in \mathbb{R}^{2\times 2}$$

Matrix Multiplication: Inner Product

- Multiply vectors, get a **scalar**

$$x^T y = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1 y_1 + x_2 y_2$$

$x \in \mathbb{R}^2 = \mathbb{R}^{2\times1}$

$x^T \in \mathbb{R}^{1\times2}$

$y \in \mathbb{R}^2 = \mathbb{R}^{2\times1}$

$x^T y \in \mathbb{R}^{1\times1} = \mathbb{R}$

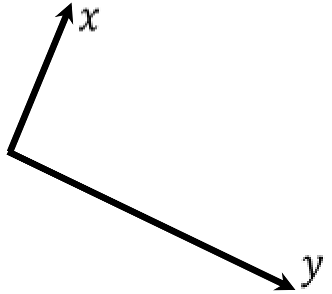
Also called **dot product**,
may write $\mathrm{x}^{\mathrm{T}}\mathrm{y} = \langle x, y \rangle$

Norms and Orthogonality

- Vectors x and y are **orthogonal** if $\langle x, y \rangle = 0$
- Norms capture "length"

Euclidean norm

$$\|x\|_2 = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^{d} x_i^2}$$



- A set of vectors $\{x_1, x_2, \dots, x_n\}$ is **orthonormal** if:
  - For all pairs x_i, x_j, we have $\langle x_i, x_j \rangle = 0$
  - For all x_i, we have $\|x_i\|_2 = 1$

Matrices May Not Commute

- In general, $AB \neq BA$

Common slip-up when
learning linear algebra

Identity Matrix

- Like "1"
- Multiplying by it gets back the same matrix or vector

- Rows & columns are the "**standard basis vectors**" e_i

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

e_1 e_2 e_n

Matrix Inverse

- If there is a B such that $AB = BA = I$
  - Then A is invertible/nonsingular, B is its **inverse**
  - Some matrices are **not** invertible!

- Notation: A^{-1}

$$\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \times \begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix} = I$$

Matrix & Vector **Operations**

Ex: feedforward neural networks. Input *x.*

- Output of layer *k* is

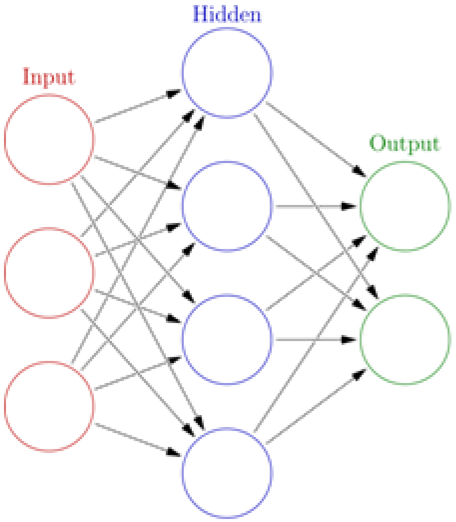
nonlinearity

$$f^{(k)}(x) = \sigma(W_k^T f^{(k-1)}(x)))$$

Output of layer k: vector

Weight **matrix** for layer k:
Note: linear transformation!

Output of layer k-1: **vector**



Wikipedia

Eigenvectors and Eigenvalues

Recall: Diagonal Matrix Multiplication

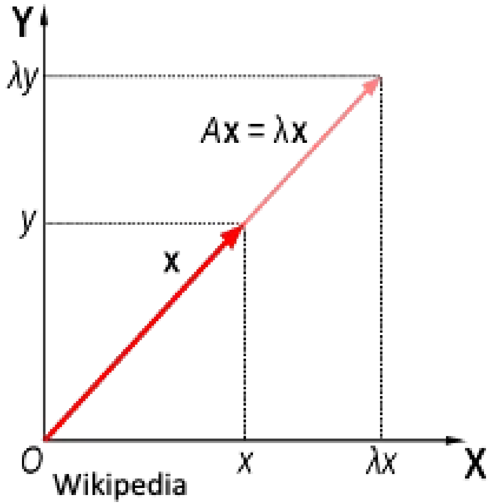
$$Dx = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} d_{11}x_1 \\ d_{22}x_2 \\ d_{33}x_3 \end{bmatrix}$$

D stretches x entry-by-entry
Different components scale differently

Eigenvectors and Eigenvalues

- Sometimes, a matrix scales a vector uniformly
- Vector $x \neq 0$ is an **eigenvector** of A if there is a scalar λ such that

$$Ax = \lambda x$$



- λ is an **eigenvalue**
- A may have many pairs $\{(\lambda_1, x_1), \dots, (\lambda_d, x_d)\}$

Eigenvalues for Identity Matrix

- Definition: (λ, x) are an eigenvalue/eigenvector pair for A if $Ax = \lambda x$
- Identity matrix: for all x, $Ix = x$

- Thus, all vectors are eigenvectors for I! (With $\lambda = 1$)

Eigenvalues for Diagonal Matrices

- Q: Which vectors satisfy $Dx = \lambda x$ for diagonal D?

$$D = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \qquad Dx = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} d_{11}x_1 \\ d_{22}x_2 \\ d_{33}x_3 \end{bmatrix}$$

- A: Standard basis vectors!

$$De_1 = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} d_{11} \\ 0 \\ 0 \end{bmatrix} = d_{11} e_1$$

Diagonals in Disguise

- Many matrices are **diagonalizable**
- We can find diagonal matrix D and orthonormal matrix Q such that

$$A = QDQ^{-1}$$

Q is orthonormal

Contain eigenvectors of A

D is diagonal

Entries are eigenvalues of A

Diagonals in Disguise

- Many matrices are **diagonalizable**
- We can find diagonal matrix D and orthonormal matrix Q such that

$$A = QDQ^{-1}$$

Orthonormal matrices:

1. Preserve length
2. Rotate & reflect

Diagonal matrices:

- Scale entry-by-entry

Further Reading

- Broad Reading: Gilbert Strang’s notes & books
  - https://math.mit.edu/~gs/LectureNotes/
  - Introduction to Linear Algebra; https://math.mit.edu/~gs/linearalgebra/ila6/indexila6.html
- Reading Ahead: math behind PCA
  - https://web.stanford.edu/class/cs168/l/l7.pdf
  - https://web.stanford.edu/class/cs168/l/l8.pdf

Linear Algebra

Basics
- Vector, Matrix
- Indexing
- Addition
- Transpose

Eigenvalues and Eigenvectors
- Satisfy $Ax = \lambda x$
- Diagonalize: $A = QDQ^{-1}$

Matrix Multiplication
- Dimensions must match
- Rule for computing
- Some special cases
- Not commutative!
- Inverse matrix
- Identity matrix
- Norm
- Orthogonal, Orthonormal

Mathematical tools for:
1. Organizing and manipulating high-dimensional data
2. Working with linear functions

Break & Quiz

- **Q 1.1**: What is $\begin{bmatrix}1 & 2\\ 3 & 1\\ 1 & 1\end{bmatrix} \times \begin{bmatrix}0\\ 1\end{bmatrix}$?
- A. $[-1\ 1\ 1]^T$
- B. $[2\ 1\ 1]^T$
- C. $[1\ 3\ 1]^T$
- D. $[1.5\ 2\ 1]^T$

Break & Quiz

- **Q 1.1**: What is $\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \end{bmatrix}$?
- A. $[-1\ 1\ 1]^T$
- **B. $[2\ 1\ 1]^T$**
- C. $[1\ 3\ 1]^T$
- D. $[1.5\ 2\ 1]^T$

Break & Quiz

- **Q 1.1**: What is $\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \end{bmatrix}$?

- A. $[-1\ 1\ 1]^T$
- **B. $[2\ 1\ 1]^T$**
- C. $[1\ 3\ 1]^T$
- D. $[1.5\ 2\ 1]^T$

Check dimensions: answer must be 3 x 1 matrix (i.e., column vector).

$$\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 * 1 + 1 * 2 \\ 0 * 3 + 1 * 1 \\ 0 * 1 + 1 * 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

Break & Quiz

- **Q 1.2**: Given matrices $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{d \times m}, C \in \mathbb{R}^{p \times n}$ What are the dimensions of BAC^T

- A. *n x p*
- B. *d x p*
- C. *d x n*
- D. Undefined

Break & Quiz

- **Q 1.2**: Given matrices $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{d \times m}, C \in \mathbb{R}^{p \times n}$

What are the dimensions of BAC^T

- A. *n x p*
- **B. *d x p***
- C. *d x n*
- D. Undefined

Break & Quiz

- **Q 1.2**: Given matrices $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{d \times m}, C \in \mathbb{R}^{p \times n}$

What are the dimensions of BAC^T

- A. *n x p*
- **B. *d x p***
- C. *d x n*
- D. Undefined

To rule out (D), check that for each pair of adjacent matrices XY, the # of columns of X = # of rows of Y

Then, B has d rows so solution must have d rows. C^T has p columns so solution has p columns.

Break & Quiz

- **Q 1.3**: A and B are matrices, neither of which is the identity. Is $AB = BA$?

- A. Never
- B. Always
- C. Sometimes

Break & Quiz

- **Q 1.3**: A and B are matrices, neither of which is the identity. Is $AB = BA$?

- A. Never
- B. Always
- **C. Sometimes**

Break & Quiz

- **Q 1.3**: A and B are matrices, neither of which is the identity. Is $AB = BA$?

- A. Never
- B. Always
- **C. Sometimes**

Matrix multiplication is not necessarily commutative.

Thanks!